Strategies in Multiagent Systems

Rodica Condurache

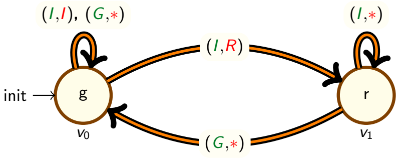
Lecture 7

Multiagent Systems

Definition

A Multiagent System is a tuple $\mathcal{M} = \langle \mathcal{AP}, \mathrm{Ag}, (\mathrm{Act}_i)_{i\in \mathrm{Ag}}, V, v_0, \tau, E\rangle$ where

- $\mathrm{Ag} = \{0, 1, ..., k\}$ is the set of agents
- Act_i is the set of actions of Agent $i \in \mathrm{Ag}$
- V is the set of states in the system
- v_0 is the initial state
- $\tau : V \to 2^{\mathcal{AP}}$ is the labeling function
- $E : V \times \mathrm{Act}_0 \times \mathrm{Act}_1 \times ... \times \mathrm{Act}_k \to V$ is the transition function



Strategies

- Given a multiagent system $\mathcal{M} = \langle \mathcal{AP}, \mathrm{Ag}, (\mathrm{Act}_i)_{i \in \mathrm{Ag}}, V, v_0, \tau, E \rangle$,
- A strategy for Agent i is a is a mapping $\sigma_i : V(\mathrm{Act}_i V)^* \to \mathrm{Act}_i$

  Agent i sees only the states and his actions!

- A strategy profile is a tuple $\bar{\sigma} = (\sigma_0, ..., \sigma_k)$ of strategies, one for each agent

- A play $\pi = v_0 v_1 v_2 ...$ is compatible with a strategy σ_i of Agent i if $\exists (a_0^0, a_1^0, ..., a_k^0)\ (a_0^1, a_1^1, ..., a_k^1)\ ...$ s.t. $\forall n \geq 0$,
  - $v_{n+1} = E(v_n, (a_0^n, a_1^n, ..., a_k^n))$ and
  - $a_i^n = \sigma_i(v_0 a_i^0 v_1 a_i^1 v_2 ... v_n)$

- $\mathrm{out}(\sigma_i)$ is the set of plays compatible with σ_i

- $\mathrm{out}(\bar{\sigma}) = \bigcap_{i \in \mathrm{Ag}} \mathrm{out}(\sigma_i)$ is the outcome of the profile $\bar{\sigma}$ - only one play!

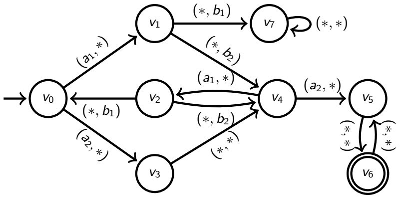
Forgetful Strategies

A strategy $\sigma_i : V(Act_i V)^* \to Act_i$ *is finite memory if*

- *there is a finite set* M *(memory states) and* $m_0 \in M$ *(initial state of memory)*
- *memory update function:* $f_i : V \times M \to M$
- *decision making function:* $g_i : V \times M \to Act_i$

During a play $\pi = v_0 v_1 \ldots$, $\sigma_i(v_0, \ldots, v_n) = g_i(v_n, m_n)$ *where* $m_n = f_i(v_n, m_{n-1})$ *for* $n \geq 1$

Example



$*$ = any action

Player 0: $\Sigma_0 = \{a_1, a_2\}$

Player 1: $\Sigma_1 = \{b_1, b_2\}$

Strategy for Player 0: in state v_4 - play first a_1 and then only a_2; in other state - always play a_2

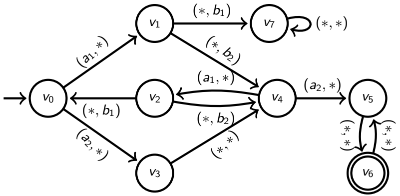
Forgetful Strategies

A strategy $\sigma_i : V(Act_i V)^* \to Act_i$ *is finite memory if*

- *there is a finite set* M *(memory states) and* $m_0 \in M$ *(initial state of memory)*
- *memory update function:* $f_i : V \times M \to M$
- *decision making function:* $g_i : V \times M \to Act_i$

During a play $\pi = v_0 v_1 \ldots$, $\sigma_i(v_0, \ldots, v_n) = g_i(v_n, m_n)$ *where* $m_n = f_i(v_n, m_{n-1})$ *for* $n \geq 1$

Example



$*$ = any action

Player 0: $\Sigma_0 = \{a_1, a_2\}$

Player 1: $\Sigma_1 = \{b_1, b_2\}$

Strategy for Player 0: in state v_4 - play first a_1 and then only a_2; in other state - always play a_2

Memory states: $M = \{f, s\}$

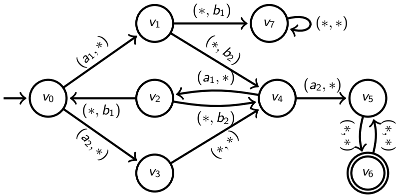
$$f_i(v, m) = \begin{cases} f & \text{if } m = f \text{ and } v \neq v_4 \\ s & \text{if } (m = f \text{ and } v = v_4) \text{ or } m = s \end{cases} \qquad g_i(v, m) = \begin{cases} a_1 & \text{if } m = f \text{ and } v = v_4 \\ a_2 & \text{if } (m = f \text{ and } v \neq v_4) \text{ or } m = s \end{cases}$$

Memoryless Strategies

A strategy $\sigma_i : V(Act_i V)^* \to Act_i$ *is memoryless iff depends only on the last state:*

$$\forall h, h' \in V^*,\ \forall v \in V,\ \sigma_i(hv) = \sigma(h'v)$$

Example



$*$ = any action

Player 0: $\Sigma_0 = \{a_1, a_2\}$

Player 1: $\Sigma_1 = \{b_1, b_2\}$

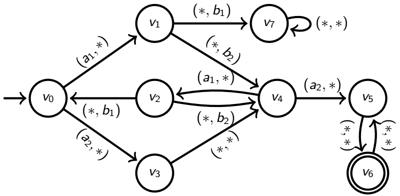
Two examples of strategies for Player 0:

- in state v_4 - play a_1; in other state - always play a_2 : $\sigma_0(v) = \begin{cases} a_1 & \text{if } v = v_4 \\ a_2 & \text{otherwise} \end{cases}$ not winning!

A strategy $\sigma_i : V(Act_i V)^* \to Act_i$ is memoryless *iff depends only on the last state:*

$$\forall h, h' \in V^*,\ \forall v \in V,\ \sigma_i(hv) = \sigma(h'v)$$

Example



$*$ = any action

Player 0: $\Sigma_0 = \{a_1, a_2\}$

Player 1: $\Sigma_1 = \{b_1, b_2\}$

Two examples of strategies for Player 0:

- in state v_4 - play a_1; in other state - always play a_2 : $\sigma_0(v) = \begin{cases} a_1 & \text{if } v = v_4 \\ a_2 & \text{otherwise} \end{cases}$ not winning!
- always play a_2 : $\sigma_0(v) = a_2,\ \forall v \in V$ winning!

Memory in Concurrent games

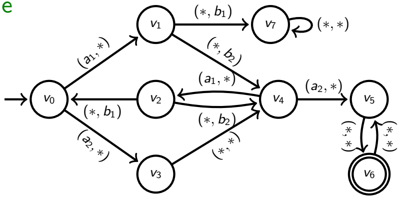
Theorem

In Rachability, Safety, Büchi and co-Büchi games $\mathcal{G}$, Player i has a memoryless winning strategy from $W_i(\mathcal{G})$.

Proof.

- see the construction of attractors in Lecture 6 □

Example



$*$ = any action

Player 0: $\Sigma_0 = \{a_1, a_2\}$
Player 1: $\Sigma_1 = \{b_1, b_2\}$

Objective of Player 0: $\textsc{Reach}(\{v_6\})$

Memory in Concurrent games

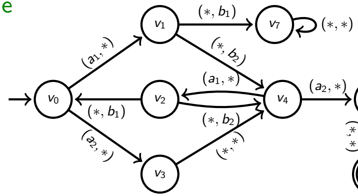
Theorem

In Rachability, Safety, Büchi and co-Büchi games $\mathcal{G}$, Player i has a memoryless winning strategy from $W_i(\mathcal{G})$.

Proof.

- see the construction of attractors in Lecture 6 □

Example



$*$ = any action

Player 0: $\Sigma_0 = \{a_1, a_2\}$
Player 1: $\Sigma_1 = \{b_1, b_2\}$

Objective of Player 0: $\text{REACH}(\{v_6\})$

Memory in Concurrent games

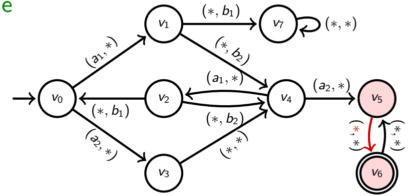
Theorem

In Rachability, Safety, Büchi and co-Büchi games $\mathcal{G}$, Player i has a memoryless winning strategy from $W_i(\mathcal{G})$.

Proof.

- see the construction of attractors in Lecture 6 □

Example



$*$ = any action

Player 0: $\Sigma_0 = \{a_1, a_2\}$
Player 1: $\Sigma_1 = \{b_1, b_2\}$

Objective of Player 0: $\textsc{Reach}(\{v_6\})$

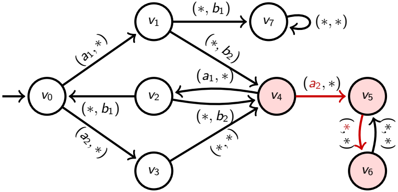
Theorem

In Rachability, Safety, Büchi and co-Büchi games $\mathcal{G}$, Player i has a memoryless winning strategy from $W_i(\mathcal{G})$.

Proof.

- see the construction of attractors in Lecture 6 □

Example



$*$ = any action

Player 0: $\Sigma_0 = \{a_1, a_2\}$
Player 1: $\Sigma_1 = \{b_1, b_2\}$

Objective of Player 0: $\textsc{Reach}(\{v_6\})$

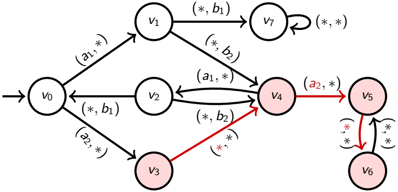
Theorem

In Rachability, Safety, Büchi and co-Büchi games $\mathcal{G}$, Player i has a memoryless winning strategy from $W_i(\mathcal{G})$.

Proof.

- see the construction of attractors in Lecture 6 □

Example



$*$ = any action

Player 0: $\Sigma_0 = \{a_1, a_2\}$
Player 1: $\Sigma_1 = \{b_1, b_2\}$

Objective of Player 0: $\textsc{Reach}(\{v_6\})$

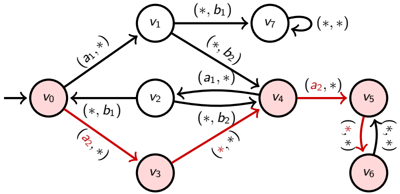
Theorem

In Rachability, Safety, Büchi and co-Büchi games $\mathcal{G}$, Player i has a memoryless winning strategy from $W_i(\mathcal{G})$.

Proof.

- see the construction of attractors in Lecture 6 □

Example



$*$ = any action

Player 0: $\Sigma_0 = \{a_1, a_2\}$
Player 1: $\Sigma_1 = \{b_1, b_2\}$

Objective of Player 0: $\textsc{Reach}(\{v_6\})$

Memory in Concurrent games

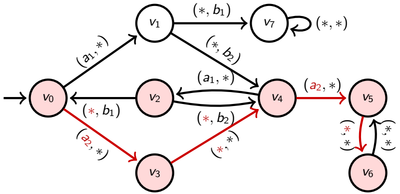
Theorem

In Rachability, Safety, Büchi and co-Büchi games $\mathcal{G}$, Player i has a memoryless winning strategy from $W_i(\mathcal{G})$.

Proof.

- see the construction of attractors in Lecture 6 □

Example



$*$ = any action

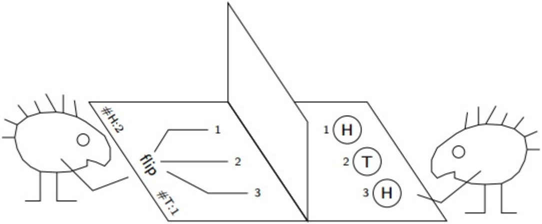
Player 0: $\Sigma_0 = \{a_1, a_2\}$
Player 1: $\Sigma_1 = \{b_1, b_2\}$

Objective of Player 0: $\textsc{Reach}(\{v_6\})$

Imperfect Information for Agents

- Imperfect information $=$ Agents don't make difference between some states
  but they know the structure of the system (game) itself

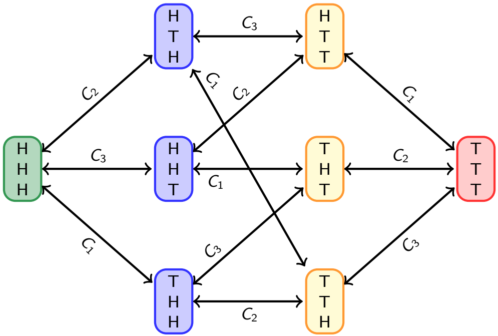
Example (3 coins Game)



Source: http://www.lsv.fr/~doyen/papers/Games_with_Imperfect_Information_Theory_Algorithms.pdf

- Left player (Player 0) knows only the number of heads and tails at each time
- Player 0 asks to flip coins depending on his information
- Right player (Player 1) obeys and flips the coins Player 0 asks for.
- Objective of Player 0: eventually *HHH* and never *TTT*

Example (3 coins Game - Model)



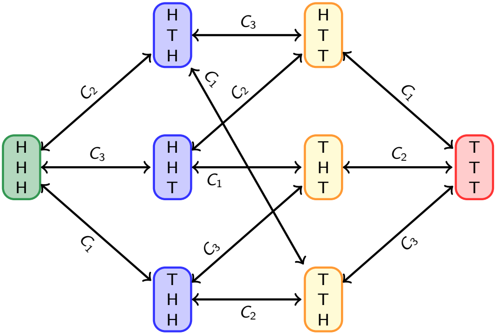
- In states with the same color, Player 0 has the same information

In general, each agent has its own observation and therefore its own coloring function!

Imperfect Information of Agents

- Imperfect information for Agent i modeled by an indistinguishability relation $\sim_i$
- Observations of Agent i = equivalence classes induced by $\sim_i$
- Let $\mathcal{O}_i$ be the set of observations for Agent i

Example



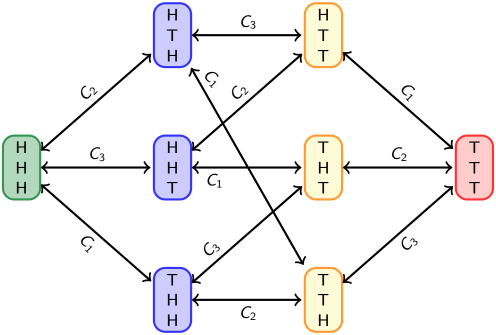
Observations for Player 0:

$$o_3 = \{hhh\}$$
$$o_2 = \{hth, hht, thh\}$$
$$o_1 = \{htt, tht, tth\}$$
$$o_0 = \{ttt\}$$
$$\mathcal{O}_0 = \{o_3, o_2, o_1, o_0\}$$

Imperfect Information of Agents

- Imperfect information for Agent i modeled by an indistinguisability relation $\sim_i$
- Observations of Agent i = equivalence classes induced by $\sim_i$
- Let $\mathcal{O}_i$ be the set of observations for Agent i

Example



Observations for Player 0:

$$o_3 = \{hhh\}$$
$$o_2 = \{hth, hht, thh\}$$
$$o_1 = \{htt, tht, tth\}$$
$$o_0 = \{ttt\}$$
$$\mathcal{O}_0 = \{o_3, o_2, o_1, o_0\}$$

Player 1 knows the exact state:

$$\mathcal{O}_1 = \{\{hhh\}, \{hth\}, \{hht\}, \ldots\}$$

Strategies with Imperfect Information

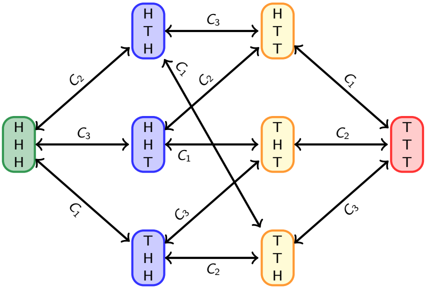
- A multiagent system with Imperfect Information is a tuple
  $\mathcal{M} = \langle \mathcal{AP}, \text{Ag}, (\text{Act}_i)_{i \in \text{Ag}}, V, v_0, \tau, E, (\mathcal{O}_i)_{i \in \text{Ag}} \rangle$,
  where $\mathcal{O}_i$ is the set of observations of Agent i (a partition of V).
- Agents play depending on the seen observations (colors)!
  - play the same action in two histories with the same "colors"
  - color of v for Agent i is $obs_i(v) = o$ s.t. $v \in o$
- A strategy for Agent i is $\sigma_i : \mathcal{O}_i(\Sigma_1 \mathcal{O}_i)^* \to \Sigma_i$
  - each agent remembers his actions and the "colors" (observations) he sees
- In games with imperfect information, objectives are required to be *observable*:

  Agent i surely wins if his objective Θ_i is satisfied by all executions with the same observations.

$$\sigma_i \textit{ is winning for Player } i \textit{ with objective } \Theta_i \subseteq (V)^\omega \textit{ iff}$$
$$\forall \pi \in out(\sigma_i), \forall \pi' \in out(\sigma_i) \textit{ s.t. } obs_i(\pi) = obs_i(\pi'), \textit{ holds } \pi' \in \Theta_i$$

Example (3 Coins Game)

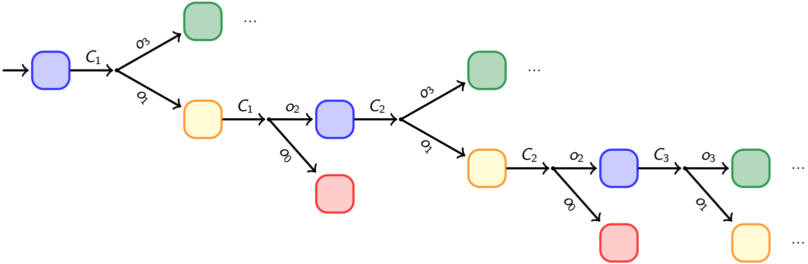
A Strategy for Player 0 in the *3 Coins Game* if the **initial state has 2 heads**:



1. flip C_1
   - If o_3 (3 heads) : keep flipping C_1 forever
   - If o_1 (1 head) : flip C_1 back and go to 2)
2. flip C_2
   - If o_3 (3 heads) : keep flipping C_1 forever
   - If o_1 (1 head) : flip C_2 back and go to 3)
3. flip C_3 forever

Example (3 Coins Game)

Graphic representation of the strategy of Agent 0 :

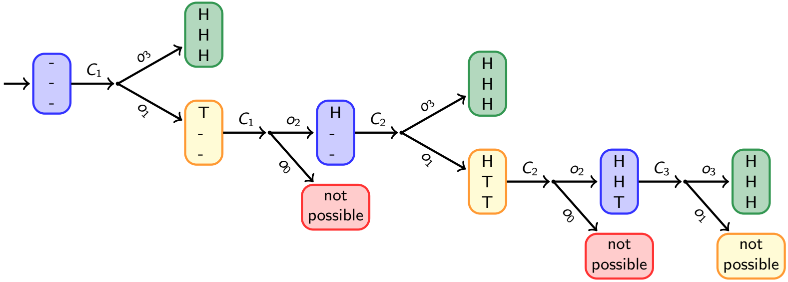


Information of Agents along executions

- Agents gain information ("learn") along executions!

Example (3 Coins Game)

Graphic representation of the strategy of Agent 0 : with information!

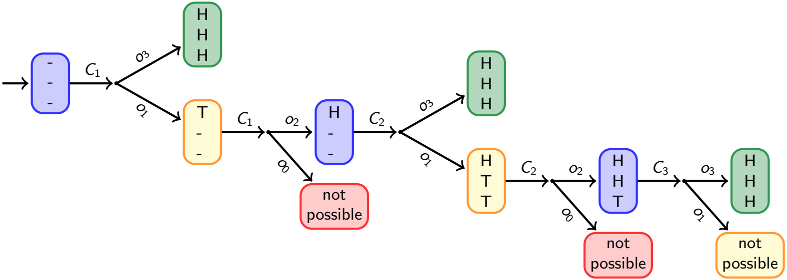


Information of Agents along executions

- Agents gain information ("learn") along executions!

Example (3 Coins Game)

Graphic representation of the strategy of Agent 0 : with information!

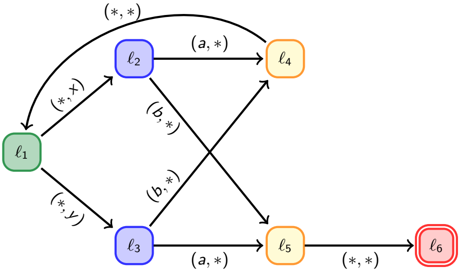


Memory is needed to win! (different actions when sees "blue" color)

In games with imperfect information, players may not have memoryless winning strategies for safety, reachability, Büchi or coBüchi objectives.

We use the information of agents when building strategies!

Example



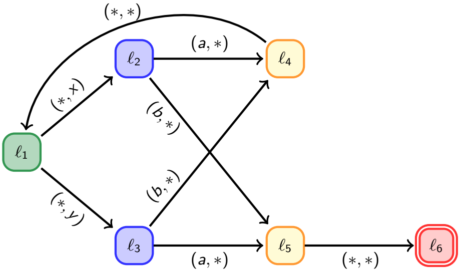
Player 0 plays : $\Sigma_0 = \{a, b\}$
Player 1 plays : $\Sigma_1 = \{x, y\}$

$\mathcal{O}_0 = \{\{\ell_1\}, \{\ell_2, \ell_3\}, \{\ell_4, \ell_5\}, \{\ell_6\}\}$

$\mathcal{O}_1 = \{\{\ell_1\}, \{\ell_2\}, \{\ell_3\}, \{\ell_4\}, \{\ell_5\}, \{\ell_6\}\}$

- Is there a strategy for Player 0 to reach ℓ_6 from ℓ_1?
- What about strategy for Player 1 to avoid ℓ_6 from ℓ_1?

Example



Player 0 plays : $\Sigma_0 = \{a, b\}$
Player 1 plays : $\Sigma_1 = \{x, y\}$

$\mathcal{O}_0 = \{\{\ell_1\}, \{\ell_2, \ell_3\}, \{\ell_4, \ell_5\}, \{\ell_6\}\}$

$\mathcal{O}_1 = \{\{\ell_1\}, \{\ell_2\}, \{\ell_3\}, \{\ell_4\}, \{\ell_5\}, \{\ell_6\}\}$

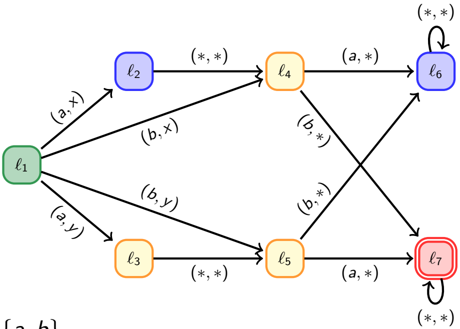
- Is there a strategy for Player 0 to reach ℓ_6 from ℓ_1? No!
- What about strategy for Player 1 to avoid ℓ_6 from ℓ_1? No!

Theorem

Reachability, Safety, Büchi and coBüchi games with imperfect information are not determined.

Exercise

Build a strategy for Player 0 to avoid ℓ_7. Compute the information sets at each step.



Player 0 plays : $\Sigma_0 = \{a, b\}$
Player 1 plays : $\Sigma_1 = \{x, y\}$

$\mathcal{O}_0 = \{\{\ell_1\}, \{\ell_2, \ell_6\}, \{\ell_3, \ell_4, \ell_5\}, \{\ell_7\}\}$

$\mathcal{O}_1 = \{\{\ell_1\}, \{\ell_2\}, \{\ell_3\}, \{\ell_4\}, \{\ell_5\}, \{\ell_6\}, \{\ell_7\}\}$

Objective of Player 0: $\Theta_0 = \text{Safe}(V \setminus \{\ell_7\})$

Safety and co-Büchi games with Imperfect Information

- When playing the game, players "compute" the information set

- In Safety games, the information sets should not contain "unsafe" states
- In coBüchi games, only a finite number of information sets can contain "final" states

Safety and co-Büchi games with Imperfect Information

- When playing the game, players "compute" the information set

- In Safety games, the information sets should not contain "unsafe" states
- In coBüchi games, only a finite number of information sets can contain "final" states

- If only Player 0 (protagonist) has imperfect information, we can reduce safety and co-Büchi games with imperfect information to games with perfect information

  Let $\mathcal{G} = \langle \mathrm{Ag} = \{0, 1\}, (\mathrm{Act}_i)_{i \in \mathrm{Ag}}, V, v_0, \tau, E, \mathcal{O}_0, \Theta_0 \rangle$ be an imperfect information game. We build the game $\mathcal{G}'$ such that:
  - the set of states : $V' = \{I \in 2^V \setminus \{\emptyset\} \mid \exists o \in \mathcal{O}_0 \ s.t. \ I \subseteq o\}$ (states are information states)
  - initial state: $I_0 = \{v_0\}$
  - transitions: $E' : V' \times \Sigma_0 \times \mathcal{O}_0 \rightarrow V'$ s.t. $E'(I, a, o) = \{v' \mid \exists v \in I \ s.t. \ v' \in E(a)\} \cap o$

Safety and co-Büchi games with Imperfect Information

- When playing the game, players "compute" the information set

- In Safety games, the information sets should not contain "unsafe" states
- In coBüchi games, only a finite number of information sets can contain "final" states

- If only Player 0 (protagonist) has imperfect information, we can reduce safety and co-Büchi games with imperfect information to games with perfect information

  Let $\mathcal{G} = \langle \mathrm{Ag} = \{0, 1\}, (\mathrm{Act}_i)_{i \in \mathrm{Ag}}, V, v_0, \tau, E, \mathcal{O}_0, \Theta_0 \rangle$ be an imperfect information game. We build the game $\mathcal{G}'$ such that:
  - the set of states : $V' = \{I \in 2^V \setminus \{\emptyset\} \mid \exists o \in \mathcal{O}_0 \ s.t.\ I \subseteq o\}$ (states are information states)
  - initial state: $I_0 = \{v_0\}$
  - transitions: $E' : V' \times \Sigma_0 \times \mathcal{O}_0 \to V'$ s.t. $E'(I, a, o) = \{v' \mid \exists v \in I \ s.t.\ v' \in E(a)\} \cap o$
  - For safety winning objective $\Theta_0 = \mathrm{Safe}(S)$: $S' = \{I \mid I \subseteq S\}$
  - For coBüchi winning objective $\Theta_0 = \mathrm{coBüchi}(T)$: $T' = \{I \mid I \cap T \neq \emptyset\}$

Safety and co-Büchi games with Imperfect Information

- When playing the game, players "compute" the information set

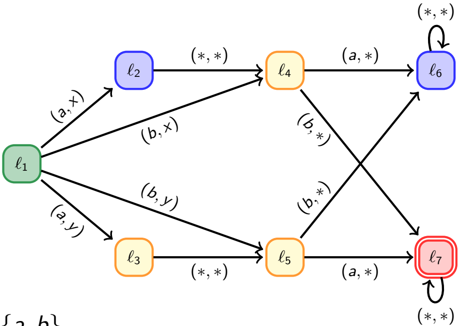
- In Safety games, the information sets should not contain "unsafe" states
- In coBüchi games, only a finite number of information sets can contain "final" states

- If only Player 0 (protagonist) has imperfect information, we can reduce safety and co-Büchi games with imperfect information to games with perfect information

  Let $\mathcal{G} = \langle \mathrm{Ag} = \{0,1\}, (\mathrm{Act}_i)_{i \in \mathrm{Ag}}, V, v_0, \tau, E, \mathcal{O}_0, \Theta_0 \rangle$ be an imperfect information game. We build the game $\mathcal{G}'$ such that:
  - the set of states : $V' = \{I \in 2^V \setminus \{\emptyset\} \mid \exists o \in \mathcal{O}_0 \; s.t. \; I \subseteq o\}$ (states are information states)
  - initial state: $I_0 = \{v_0\}$
  - transitions: $E' : V' \times \Sigma_0 \times \mathcal{O}_0 \to V'$ s.t. $E'(I, a, o) = \{v' \mid \exists v \in I \; s.t. \; v' \in E(a)\} \cap o$
  - For safety winning objective $\Theta_0 = \mathrm{Safe}(S)$: $S' = \{I \mid I \subseteq S\}$
  - For coBüchi winning objective $\Theta_0 = \mathrm{coBüchi}(T)$: $T' = \{I \mid I \cap T \neq \emptyset\}$
- Büchi games with imperfect information are more difficult to solve!

Exercise

Build the perfect information game corresponding to the below game where Player 0 has imperfect information and safety objective to avoid ℓ_7. Solve the resulting game.



Player 0 plays : $\Sigma_0 = \{a, b\}$
Player 1 plays : $\Sigma_1 = \{x, y\}$

$\mathcal{O}_0 = \{\{\ell_1\}, \{\ell_2, \ell_6\}, \{\ell_3, \ell_4, \ell_5\}, \{\ell_7\}\}$

$\mathcal{O}_1 = \{\{\ell_1\}, \{\ell_2\}, \{\ell_3\}, \{\ell_4\}, \{\ell_5\}, \{\ell_6\}, \{\ell_7\}\}$

Objective of Player 0: $\Theta_0 = \mathrm{Safe}(V \setminus \{\ell_7\})$

Bibliography

- http://www.lsv.ens-cachan.fr/~dwb/gtc.pdf
- https://www.irif.fr/~serre/Files/MPRI-notes-2015.pdf
- http://www.lsv.fr/~doyen/papers/Games_with_Imperfect_Information_Theory_Algorithms.pdf